Joint Centre for Mesoscale Meteorology, Reading, UK

Fluctuations in an equilibrium convective ensemble
Parts I and II: Theoretical formulation and Experimental validation

JCMM Internal Report No 137

Forecasting ResearchTechnical Report 396

September 13, 2002

Brenda G. Cohen

George C. Craig

Met Office Joint Centre for Mesoscale Meteorology Department of Meteorology
University of Reading PO Box 243 Reading RG6 6BB United Kingdom
Tel: +44 (0)118 378 8425 Fax: +44 (0)118 378 8791
www.metoffice.com

Designed and produced by the Met Office © Crown copyright 2002 02/0392
Met Office and the Met Office logo are registered trademarks

Fluctuations in an equilibrium convective ensemble

Parts I and II: Theoretical formulation and Experimental validation

BRENDA G. COHEN [1] AND GEORGE C. CRAIG

Department of Meteorology, University of Reading, UK

September 13, 2002

[1]Corresponding author address: Brenda G. Cohen, Department of Meteorology, University of Reading, PO Box 243, Reading, RG6 6BB, UK. Email: B.G.Cohen@reading.ac.uk

ABSTRACT

The equilibrium fluctuations of a field of cumulus clouds under homogeneous large-scale forcing are derived statistically, by exploiting the analogy between convection and an ideal gas. In the limit of non-interacting convective cells, the statistics of these convective fluctuations can be written in terms of the large-scale, externally constrained properties of the system. Using this framework, the probability density function of individual cloud mass fluxes is shown to be exponential. An analytical expression for the distribution function of total mass flux over a region of given size is also derived, and the variance of this distribution is found to be inversely related to the mean number of clouds in the ensemble. In a companion paper, these theoretical predictions are tested experimentally against Cloud Resolving Model data.

The theoretical predictions derived in Part I are tested against 3-dimensional Cloud Resolving Model (CRM) simulations of radiative-convective equilibrium. The exponential distribution of individual mass fluxes obtained in these experiments is shown to be in direct agreement with theory. The distribution function for total mass flux over a finite region agrees well with the theoretical prediction, but only after a rescaling of the variance to account for spatial regularity in the simulated cloud fields. This regularity is associated with a short-range inhibition, which has greater influence when the convective forcing is strong and the clouds closer together, suggesting that the theory will only be correct in the limit of weakly forced convection. In practice, the mass flux distribution function for any sized region can be obtained from the theoretical formula by applying a rescaling factor to the variance that is only a function of the forcing.

Fluctuations in an equilibrium convective ensemble
Part I: Theoretical formulation

1 Introduction

The parameterisation of cumulus convection in numerical models of the atmosphere is usually based upon a local equilibrium hypothesis, where it is assumed that the average properties of the (unresolved) convection within each grid-box can be determined solely in terms of large-scale (resolved) conditions. In particular, the heating, moistening and momentum sources of the convection are functions of the temperature, moisture content and other model variables in the relevant column (Gregory and Miller, 1989; Tiedke, 1989; Emanuel, 1991). This assumption can only be exactly true in the limit of averaging over an infinite number of clouds. For a region of finite size, the average convective properties will fluctuate around the equilibrium average, depending on how many clouds happen to be in the region at any given moment.

Deviations from equilibrium are important for at least three reasons. First, subgrid-scale convective fluctuations are a significant source of variability for the larger-scale climate (Lin and Neelin, 2000). Secondly, quantitative precipitation forecasting often requires predictions for very small regions (*e.g.* river catchments) which are only probabilistically related to the larger-scale flow that is observed and forecast (Fritsch *et al.*, 1998). Finally, convective variability is a significant source of model error in deterministic numerical weather prediction models and can influence the spread of ensemble forecasts (Buizza *et al.*, 1999; Palmer, 2001).

It is therefore an important task to characterise convective variability. In particular,

one would hope to predict the statistics of the convective fluctuations in terms of the same large-scale properties that determine the mean behaviour. This paper attempts to provide such a characterisation using standard methods from physics for describing fluctuations about an equilibrium. A companion paper describes tests of the theoretical predictions using a Cloud Resolving Model.

Before attempting an equilibrium fluctuation theory, it is important to distinguish the convective variability due to local fluctuations in an equilibrium situation from variability that occurs when the convection is not in equilibrium at all. Where convection is spontaneously organised into mesoscale structures such a squall lines, there may be up-scale organisation into convective clusters and superclusters, in which case there is no distinction between the convective system and the evolving synoptic flow (Chen *et al.*, 1996; Moncrieff and Klinker, 1997). In this case there will be no deterministic relationship between the large-scale and the statistics of the convective variability, because the convective fluctuations (individual squall lines) will significantly influence these larger scales. The discussion in this paper will be restricted to disorganised convection in an unsheared environment, a situation where there is a well-defined scale separation between the large-scale flow and the convective cloud spacing. This implies that the convection will be in equilibrium with its environment if averaged over a sufficiently large area. The aim is then to characterise the variability within a smaller sub-region, such a river catchment or numerical model grid-box. The extent to which the arguments can be applied to organised convection will be discussed at the end of the paper.

The statistical characterisation of small-scale variability in turbulent flows can be a very complex task, even for the simple case of homogeneous turbulence (Lesieur, 1997).

But moist atmospheric convection has a useful simplification that is not found in most turbulent systems. The fraction of the atmosphere occupied by active convective motions is small (Cotton and Anthes, 1989) and, in the absence of convective organisation, it may be a reasonable hypothesis to regard convection as composed of isolated entities that interact with each other only through the mean flow. This is precisely analogous to the 'ideal' assumption for an ideal gas, where the gas is presumed to be composed of particles whose statistical properties are determined by the energy and particle number of the whole system, and not influenced significantly by the details of how the molecules interact locally in collisions (Landau and Lifshitz, 1968). Of course, any real gas shows deviations from ideal behaviour as the density becomes larger and, similarly, we expect that in some circumstances local interactions of convective clouds through cold pools or compensating subsidence may be important. However, the ideal gas hypothesis has been shown to be reasonable for many situations of practical interest, and it may be hoped that a characterisation of ideal convective behaviour will also prove useful. Based on this picture, a statistical model will be developed, along the lines of Cho (1978) or Randall and Huffman (1980), but utilising the assumptions of the underlying physical model.This systematic approach provides a firm foundation for precise quantitative testing of the theory and extension to the non-ideal world of organised convection.

This paper will therefore exploit the analogy with ideal gas theory in developing formulae for convective statistics. Nevertheless, since there are significant differences between the underlying conceptual models of the gas and cloud systems, the mathematical derivations presented here will be motivated by the convective problem, rather than following the standard textbook approach to statistical mechanics. Section 2 discusses

the assumptions underlying the development in more detail, and attempts to identify the physical parameters that will form the basis of the theory. The main results are derived in sections 3 and 4, which present probability density functions for the mass flux of a single cloud and the total mass flux over a region of given size, respectively. Section 5 presents a discussion of the implications of the results, the extension to organised convection, and methods for testing the predictions.

2 Assumptions and constraints

a. The "ideal" equilibrium convecting atmosphere

In this section, the assumptions underlying the statistical derivations of the subsequent sections are discussed. An attempt will be made to provide physical motivation for each statement, but these cannot be regarded as compelling arguments. Indeed, other self-consistent theories could be constructed using different sets of assumptions. The particular choices made here can only be assessed by testing the conclusions derived from them.

1) Large-scale constraints on the convective ensemble

We take as a starting point a conventional 'mass flux' picture of the relationship between an ensemble of convective clouds and its environment (Ooyama, 1971; Gregory, 1997). The atmosphere is rendered unstable to moist convection on the large-scale through some combination of radiative cooling, large-scale ascent that produces adiabatic cooling, and surface fluxes of heat and moisture. The convection, consisting of moist updraughts and downdraughts, acts to reduce the instability by a net transport of mass

upwards. Some or all of this mass descends between the convective updraughts, warming the atmosphere by adiabatic compression, and thus tending to restore stability. If the area occupied by the convective up and downdraughts is small and re-evaporation of precipitation is ignored, it can be shown that the net convective mass flux at equilibrium is constrained by the need to produce subsidence warming equal to the large-scale cooling that forces the convection (Gregory and Miller, 1989). We therefore assume that the mean convective mass flux over the area of interest at some arbitrary height, $\langle M \rangle$ (in units of kgs^{-1}) is constrained by the large-scale flow. The angle brackets denote an ensemble mean over all possible realisations of the convection consistent with the large-scale forcing conditions. The calculation of the mean convective mass flux at cloud-base level is a standard part of 'mass flux' cumulus parameterisations.

The variability of equilibrium convection results from the fact that the mean mass flux, $\langle M \rangle$, is divided up among a number of finite-sized clouds. A given region may contain various numbers of clouds of various sizes, although the total mass flux must approach the average, $\langle M \rangle$, if the region is very large. It is thus necessary to characterise the degree to which the convection is broken up, with some measure of the typical number of clouds or their typical size. The size of a convective cloud is likely to be determined by properties local to that cloud – the initial perturbation that triggers it, and the mixing processes that accompany its development. We therefore take as a second fundamental parameter describing the convective ensemble the mean mass flux per cloud, $\langle m \rangle$ (kgs^{-1}). Some evidence for this is provided by the cumulus ensemble simulations of Robe and Emanuel (1996) that show the convective vertical velocity to be almost independent of the large-scale forcing and mean convective mass flux. This invariance of the mass flux

to the forcing has also been noted by Cohen (2001).

2) Scale separation

We assume that a scale separation exists in time and space between variations in the large-scale forcing and in the convective ensemble. In particular, it is assumed that it is possible to define a region that is small enough in space that the forcing of convection does not vary significantly over the region and yet large enough to contain many clouds (Yanai *et al.*, 1973; Arakawa and Schubert, 1974). This assumption implies that the average quantities, $\langle M \rangle$ and $\langle m \rangle$, which characterise the convective ensemble, are well-defined. We further assume that the forcing varies sufficiently slowly in time that the convection has time to reach and remain close to equilibrium. This implies that $\langle M \rangle$ and $\langle m \rangle$ can be determined solely in terms of the large-scale forcing.

3) Weak interactions

If the convective clouds are sufficiently separated it may be reasonable to ignore the interactions of a cloud with its immediate neighbours. The cloud will feel only an aggregate change to its environment resulting from the effects of all the other clouds. In other words, the clouds will interact only through their effects on the large-scale state of the atmosphere. This idealisation of the convective behaviour implies that any aspect of the convective variability can be described in terms of the mean properties of the ensemble, *i.e.* $\langle M \rangle$ and $\langle m \rangle$.

4) Equal a priori probabilities

A final assumption, key to all of equilibrium statistical mechanics, is the postulate of equal *a priori* probabilities. We assume that clouds are equally likely to occur in any

location and with any mass flux, subject only to the constraints described above, namely that the mean mass flux, $\langle M \rangle$, and the mean mass flux per cloud, $\langle m \rangle$, are as required by the large-scale forcing. This assumption cannot be justified, except by its success in the statistical description of many natural phenomena.

b. The analogy with the ideal gas

Before developing the consequences of the above assumptions, it is interesting to compare this conceptual model of convective activity with the ideal gas model that motivated it. An ideal gas is assumed to be a system of particles that satisfy the following assumptions. Firstly, there is a scale separation that allows one to consider a large number of particles with well defined averages. Additionally, the gas system is assumed to be in equilibrium, and the particles are assumed to be weakly interacting, so that their statistics are functions of the mean properties of the system. Finally, all states of the system consistent with these assumptions are assumed to be equally probable. Real gases tend to show ideal behaviour only in the low-density limit; when particles are closely packed, interactions become significant and non-ideal behaviour is observed.

The large-scale constraints that the system of particles must satisfy are to some extent analogous to those of the convective system. In particular, the requirement that energy, E, is conserved (or conserved in the mean for a system in contact with a heat bath) corresponds to the requirement that convective mass flux, $\langle M \rangle$, is constrained in the mean. The temperature, T, of the gas is related to the mean energy per degree of freedom for the system, and is thus analogous to the mean mass flux per cloud, $\langle m \rangle$. A measure of the disorder of the system is given by the entropy, which can be defined as $S = \frac{\partial E}{\partial T}$. The corresponding quantity for the convective system is $\frac{\partial \langle M \rangle}{\partial \langle m \rangle}$, equal to $\frac{\langle M \rangle}{\langle m \rangle}$

if $\langle m \rangle$ is independent of the large-scale forcing. This can be understood physically by noting that $\langle N \rangle \equiv \frac{\langle M \rangle}{\langle m \rangle}$ is the mean number of clouds in the region. A larger entropy, or more disordered state, occurs when the mean mass flux, $\langle M \rangle$, is divided up among a large number of small clouds, rather than a small number of large clouds. Note that $\langle M \rangle$ and $\langle N \rangle$ are extensive variables, like E and S in the gas system, and are thus proportional to the size of the system, while $\langle m \rangle$, like the temperature of the gas, is an intensive quantity, independent of system size.

There are, however, several aspects of the ideal gas that are not reflected in the convection problem. Firstly, the number of particles in a gas is fixed, while there is no similar constraint on the number of clouds. In this sense the problem is more analogous to a photon gas, than a gas composed of atoms or molecules (Chandler, 1987). Furthermore for an ideal gas, energy is required to reduce the volume (pressure work). There is no analogous term in the convective problem described here. Finally, there is, of course, no quantisation of convective mass flux.

3 Probability distribution of mass flux per cloud

In this section, the distribution of individual cloud mass fluxes is described in terms of the mean mass flux per cloud, $\langle m \rangle$, which is assumed to be externally constrained. Since we assume that the clouds are non-interacting and thus independent of their neighbours, we do not need to consider the spatial locations of the clouds in this calculation. The derivation here is formulated in terms of stochastic processes, but the resulting exponential probability distribution of cloud mass fluxes is precisely analogous to the Boltzmann distribution of molecular energies for an ideal gas, obtained from the statistical mechanics

of the canonical ensemble.

To compute how the mass flux in a given region is distributed among the individual clouds, we consider the following construction. The total mass flux is represented as a line, with the mass fluxes of the individual clouds corresponding to intervals on the line (Figure 1). The fixed mean mass flux per cloud can be written equivalently as a constraint on the mean number of clouds *per unit mass flux*:

$$\lambda = \frac{\langle N \rangle}{\langle M \rangle} \equiv \frac{1}{\langle m \rangle}, \tag{1}$$

or equivalently the number of dividing points in a given length. Since the mass fluxes of the individual clouds are assumed to be uncorrelated, the number of clouds contributing to a given mass flux (i.e. the number of points in a given length) is a random variable, whose distribution, $p_N(n)$, can be described by the Poisson point process:

$$p_N(n) \equiv Pr\{N((0, \langle M \rangle]) = n\} = \frac{(\lambda\langle M \rangle)^n exp^{-\lambda\langle M \rangle}}{n!} \quad \text{for } n = 0, 1, ..., \tag{2}$$

with 'rate' parameter $\lambda = \frac{1}{\langle m \rangle}$.

It is straightforward to show from Equation (2) that the distribution of interval lengths between points is governed by an exponential distribution, and hence that the mass flux of individual clouds is exponentially distributed. This result can be demonstrated simply for a mass flux interval containing exactly two clouds, but can easily be extended to a longer interval containing any number of clouds.

Consider an interval of mass flux, sub-divided into two clouds with individual mass fluxes in the range m_1 to $m_1 + dm_1$ and m_2 to $m_2 + dm_2$, as shown in Figure 1. The joint probability density function for M_1 and M_2 can be written as a product of the probabilities of exactly one event occurring in each of the intervals $(m_1, m_1 + dm_1]$ and

$(m_2, m_2 + dm_2]$ (Taylor and Karlin, 1994), *i.e.*:

$$\begin{aligned} p_{M_1,M_2}(m_1,m_2)dm_1dm_2 &= Pr\{N((0,m_1]) = 0\} \\ &\times Pr\{N((m_1,m_1+dm_1]) = 1\} \\ &\times Pr\{N((m_1+dm_1,m_1+dm_1+m_2]) = 0\} \\ &\times Pr\{N((m_1+dm_1+m_2,m_1+dm_1+m_2+dm_2]) = 1\}. \end{aligned} \tag{3}$$

Using Equation (2), this can be re-written as:

$$p_{M_1,M_2}(m_1,m_2)dm_1dm_2 = (e^{-\lambda m_1})(e^{-\lambda m_2})(\lambda dm_1 e^{-\lambda dm_1})(\lambda dm_2 e^{-\lambda dm_2}). \tag{4}$$

Upon dividing both sides of (4) by $(dm_1)(dm_2)$, and taking the limit as $dm_1, dm_2 \to 0$, it is found that:

$$p_{M_1,M_2}(m_1,m_2) = (\lambda e^{-\lambda m_1})(\lambda e^{-\lambda dm_2}). \tag{5}$$

Finally, using the fact that individual clouds are independent, the probability distribution of mass flux per cloud can be written in general form as:

$$p(m)dm = \frac{1}{\langle m \rangle} e^{-m/\langle m \rangle} dm. \tag{6}$$

This result is analogous to the Boltzmann distribution, with $\langle m \rangle$ playing the role of temperature.

Since the mean number of clouds contributing to the total mass flux is $\langle N \rangle$, the average number of clouds within a mass flux interval (or equivalently, within a given spatial region) possessing a mass flux between m and $m + dm$ is:

$$\begin{aligned} d\overline{n}(m) &= \langle N \rangle p(m)dm \\ &= \frac{\langle N \rangle}{\langle m \rangle} e^{-m/\langle m \rangle} dm. \end{aligned} \tag{7}$$

This final equation is in a form suitable for experimental verification.

4 Probability distribution of total mass flux

a. Probability density function

The probability distribution of the total mass flux of the cloud ensemble in a given region of finite size, $p(M)$, can be derived by combining the distribution of individual mass fluxes from (6) with the Poisson distribution of cloud number per unit mass flux from (2).

It is most straightforward to consider firstly the situation where the number of clouds in the ensemble is fixed to be exactly N. In this situation, the probability density function for the total mass flux of the ensemble, $p^{(N)}(M)$, can be calculated as an N-fold convolution over $p(m)$ (Taylor and Karlin, 1994); *i.e.*

$$p^{(N)}(M) = \int_0^M p^{(N-1)}(M-u)p(u)du. \tag{8}$$

This gives, on recursive substitution of (6) into (8):

$$p^{(N)}(M) = \frac{1}{(N-1)!}\left(\frac{1}{\langle m\rangle}\right)^N M^{N-1}e^{-M/\langle m\rangle}. \tag{9}$$

Next, in order to relax the assumption of fixed N, we sum over (9) for all values of $N = n$, weighted by the probability of there being exactly n clouds, $p_N(n)$, *i.e.*

$$p(M) = \begin{cases} \sum_{n=1}^{\infty} p^{(n)}(M)p_N(n) & \text{for } n \geq 1 \\ p_N(n) \equiv e^{-\langle N\rangle} & \text{for } n = 0. \end{cases} \tag{10}$$

$p(M)$ has a discrete component, corresponding to $n = 0$, and a continuous distribution when $n \geq 1$.

For $n \geq 1$, then, we find by substituting Equations (9) and (2) into (10):

$$p(M) = \frac{\langle N\rangle}{\langle m\rangle}e^{-\langle N\rangle}e^{-M/\langle m\rangle}\sum_{n=1}^{\infty}\frac{\left(\frac{\langle N\rangle}{\langle m\rangle}M\right)^{n-1}}{n!(n-1)!}. \tag{11}$$

The summation in Equation (11) is of standard form, and can be written as:

$$\sum_{n=1}^{\infty} \frac{x^{n-1}}{n!(n-1)!} \equiv \frac{I_1(2\sqrt{x})}{\sqrt{x}}, \tag{12}$$

where $I_1(x)$ is the modified Bessel function of order 1.

Hence, for our cloud system, the probability distribution of total mass flux is given by:

$$p(M) = \left(\frac{\langle N \rangle}{\langle m \rangle}\right)^{1/2} e^{-\langle N \rangle} M^{-1/2} e^{-M/\langle m \rangle} I_1\left(2\sqrt{\tfrac{\langle N \rangle}{\langle m \rangle} M}\right). \tag{13}$$

This expression is valid for a convective ensemble containing any number of clouds (greater than zero). As discussed further in Section 5, a version of this formula has been derived in a related context by Rodriguez-Iturbe *et al.* (1987).

The form of this distribution is not easily recognisable, and so the shape of $p(M)$ is plotted in Figures 2(a) and 2(b) for two different values of $\langle N \rangle$, corresponding to a small region ($\langle N \rangle = 5$) and a region containing a large number of clouds ($\langle N \rangle = 68$). The distribution becomes more symmetrical around its mean value as $\langle N \rangle$ increases, and the variance decreases correspondingly.

b. Fluctuations in total mass flux

The variance of the total convective mass flux in a region can be calculated from the distribution calculated in the previous section; however, it can also be obtained directly from the initial assumptions in a way that gives some physical insight into the factors controlling the variability. The total mass flux of an arbitrary n-cloud system can be written as:

$$M(n) = \sum_{i=1}^{n} m_i, \tag{14}$$

where m_i is the mass flux of an individual cloud. Here, the probability distribution of each m_i is identical (given by Equation (6)), and the clouds are assumed to evolve independently of one another. Now, defining the probability distribution of the number of particles in the general case as $p(n)$, the theory of *random sums* can be invoked to provide predictions for the mean and variance of $M(n)$ in terms of $\langle m \rangle$ and $\langle N \rangle$ (the derivation is straightforward and is given in Taylor and Karlin, 1994). Specifically, for the assumed random, Poisson distribution of cloud number (Equation (2)), the predictions are:

$$\langle M \rangle = \langle N \rangle \langle m \rangle, \tag{15}$$

in agreement with Equation (1), and

$$\langle (\delta M)^2 \rangle = 2 \langle N \rangle \langle m \rangle^2. \tag{16}$$

Equation (16) can be normalised by the mean mass flux, to give:

$$\frac{\langle (\delta M)^2 \rangle}{\langle M \rangle^2} = \frac{2}{\langle N \rangle}. \tag{17}$$

This final expression is simply a mathematical statement of the intuitive result that the variance scales inversely with the number of clouds in the ensemble.

5 Discussion

a. Meaning of parameters

Expressions have been derived in this paper for the distribution of mass flux carried by individual clouds and of the total convective mass flux over a region of finite size, for a convecting atmosphere in equilibrium with the large-scale forcing. Two parameters are

required, the mean convective mass flux, $\langle M \rangle$, and the mean mass flux per cloud, $\langle m \rangle$. As noted previously, $\langle M \rangle$ is determined by the requirement that the convection balance the large scale forcing when averaged over a large region, and is provided, for example, by the closure in a mass flux convection scheme.

The second parameter, $\langle m \rangle$, is not necessarily a function of the large-scale forcing. Some results (Robe and Emanuel, 1996; Cohen, 2001) suggest that it may even be independent of the forcing; that is, that the response to a change in forcing is to change the number of clouds, rather than the properties of the individual clouds. This in turn indicates that the mass flux of an individual cloud is related to the initial perturbation that triggers it and the entrainment processes that modify its mass flux over time. At present there can be little certainty about how $\langle m \rangle$ is determined since the crucial processes are not observed in sufficient detail, nor resolved in most numerical simulations. However, the key role that this parameter plays in determining the overall convective variability makes it an attractive candidate for future research.

b. Interpretation of mass flux variability

The variability of convective mass flux in a given region will be small if the number of clouds is large, *i.e.* if the size of the region is large in comparison to the typical cloud separation distance. Thus, a necessary condition for convection in a finite region, such as a grid box in a numerical model, to be in equilibrium with the average forcing over that region is that the region size Δx be larger than the mean cloud spacing, $L = (A/\langle N \rangle)^{1/2} = (\langle m \rangle A/\langle M \rangle)^{1/2}$, where A is the size of area containing on average $\langle N \rangle$ clouds. It is interesting to note that, since $\langle m \rangle$ appears to be independent of the forcing, L is inversely related to $\langle M \rangle$. Where convection is strongly forced, there are

more clouds to average over and less convective variability in a region of a given size. An equilibrium convective parameterisation will be most accurate where it is most needed.

An estimate of the magnitude of convective variability can easily be obtained from typical values of $\langle m \rangle$ and $\langle M \rangle$. For a purely radiative-convective equilibrium, latent heat release must balance radiative cooling S. If the rate of latent heating is estimated crudely as the convective mass flux multiplied by a typical water vapour mass mixing ratio, q, we obtain

$$l_v q \frac{\langle M \rangle}{A} = S, \tag{18}$$

where l_v is the latent heat of condensation. Values of S =250 Wm^{-2} and q =10 gkg^{-1} give $\langle M \rangle / A = 10^{-2}\ kgm^{-2}s^{-1}$. Estimating $\langle m \rangle$ for a deep convective cloud with a 1 ms^{-1} vertical velocity and a size of 1 km^2 gives $\langle m \rangle = 10^6\ kgs^{-1}$. The characteristic cloud spacing is then L =10 km. The normalised standard deviation of convective variability in a square region with side Δx is from (Equation 17)

$$\frac{\delta M}{M} = 2^{1/2} \frac{L}{\Delta x}, \tag{19}$$

which gives a value of only 0.14 if $\Delta x = 100\ km$, a typical resolution for a global model, but 0.7 for $\Delta x = 20\ km$, a typical mesoscale model resolution. In regions of substantial mean ascent the convective forcing will be larger and the variability smaller: for example an increase in forcing by a factor of four will reduce the variability to half the radiative-convective value. By comparison, in the stochastic parameterisation scheme of Buizza *et al.* (1999), the optimum value of random perturbation applied to the equilibrium tendency was found to be in the range 0.5 to 1.5, for a $1^\circ \times 1^\circ$ region. This introduces a magnitude of variability that is larger than would be predicted from the statistical theory,

but is perhaps reasonable for situations in which convective organisation is prevalent, as discussed below.

c. Extension to mesoscale convective organisation

The extension of the results here to organised convection may be trivial, or may be impossible. Much convection, even in areas of relatively homogeneous forcing such as the tropical oceans is organised into mesoscale convective systems (*e.g.* Rickenbach and Rutledge, 1998). This organisation is presumably associated not with mesoscale variability in the convective environment, but with local triggering of convection by cold pools which are strong in the presence of environmental vertical wind shear and strong downdraughts. The obvious way to extend the present theory to this situation is to assume that, although the individual clouds are clustered into mesoscale systems, the systems themselves are randomly located. In this case, one can simply interpret $\langle m \rangle$ as the mean mass flux per convective system and use the same expressions for the convective variability. Since $\langle m \rangle$ will be larger for a mesoscale convective system than for a single cloud, there will be fewer of these units in a given area and hence a larger variability.

There are, however, problems with this simple generalisation that stem from two basic assumptions of the theory. Firstly, the area occupied by an individual cloud (or convective system) has been ignored; an assumption that is only valid if the fractional area occupied by the cloud (or system) is small. Since mesoscale convective systems often contain extensive regions of stratiform precipitation that may account for a significant proportion of the overall mass flux, the theory may only be relevant for rather large regions. Secondly, the predictions assume that the forcing of convection is homogeneous over a region that is large in comparison to the convective cloud spacing so that quantities

such as $\langle M \rangle$ are well defined. Since the spacing of mesoscale convective systems will be hundreds of kilometres, this scale separation may be only rarely satisfied. Moreover, in the event of spontaneous up-scale organisation, as has been seen in 2D Cloud Resolving Model experiments (Peng *et al.*, 2001; Grabowski and Moncrieff, 2001), scale separation may no longer be a meaningful concept.

d. Experimental validation

A comparison of the distributions found in this work with observations is essential, but difficult for two reasons. First, sufficient data is needed to define the mass flux distribution, especially at the large end of the distribution where the most important events lie and various alternative distribution functions show their largest differences. This is particularly difficult since, as indicated above, the parameters of the distribution function depend upon the forcing of convection, so that data must be stratified or rescaled according to the environmental conditions. Secondly, the constrained variable in the theory, convective mass flux, is difficult to observe directly. Comparisons of the predictions for mass flux to observations of other quantities such as radar reflectivity or visible cloud size are unlikely to be quantitative, and may even show qualitative differences due to the complex interactions of dynamics and microphysics that produce the observed patterns. The existing observations will be reviewed briefly below. An alternative way of testing the theory is to use simulated data from a cloud resolving model; this is the subject of Part II of this work.

Direct observations of the mass flux of cumulus clouds have been made by Jorgensen and Lemone (1989), although their data was not sufficiently representative to determine if the distribution was exponential. Exponential distributions of cloud size using im-

agery from aircraft or satellite have been reported by a number of authors (Plank, 1969; Hozumi *et al.*, 1982; Astin and Latter, 1998), while in other studies that are based on radar reflectivity, data have been fitted to a log-normal distribution (Lopez, 1978; Lopez *et al.*, 1984). In this case, however, the difference between log-normal and exponential distributions may be unclear, since reflectivity will not reflect the population of small clouds that contain few precipitation particles. Finally, many analyses, particularly more recent work, have shown power law distributions of cloud size, usually with a scale break (*e.g.* Cahalan and Joseph, 1989; Sengupta *et al.*, 1990; Nair *et al.*, 1998). This may result from non-ideal behaviour in the mass flux, but may also result from the different processes that affect cloud size. In particular, clouds persist and spread even after the convective updraught weakens, and are thus more likely than the updraughts themselves to interact and merge with their neighbours. This merging may lead to percolation and the emergence of power-law scaling (Stauffer and Ahorony, 1992). It is interesting that some of the 'young' cloud fields studied by Sengupta *et al.* (1990) and Nair *et al.* (1998) show distributions that are not far from exponential, while the more mature cloud fields are clearly fit by a power law. It may be that the visible cloud size more closely represents the mass flux of a cloud early in its life cycle.

The spatial distribution of clouds, assumed to be random in this paper, and the corresponding distribution function for mass flux can also be compared to observations. The results here are inconclusive since random distributions (Nair *et al.*, 1998), spatial regularity (Ramirez *et al.*, 1990; Nair *et al.*, 1998), and clustering of clouds (Cahalan and Joseph, 1989; Sengupta *et al.*, 1990) have all been observed. The behaviour is presumably sensitive to the environment where the clouds are forming, but it is not

clear what specific factors are responsible for a particular behaviour.

A better proxy for mass flux than cloud area might be precipitation, which is directly related to the column-integrated latent heating. The spatial distribution of precipitation has been extensively studied in the hydrological literature (Gupta and Waymire, 1979; Kavvas and Delleur, 1981; Eagleson, 1978), and one study in particular is especially relevant to the work presented here. Rodriguez-Iturbe *et al.* (1987) obtained a theoretical cumulative rainfall probability distribution exactly analogous to Equation (13), and compared this prediction against monthly precipitation data from Denver, Colorado, collected over a time-span of 28 years. The agreement was poor. However, this comparison assumed constant values of $\langle M \rangle$ and $\langle N \rangle$ throughout the observing periods, whereas in reality these quantities vary significantly as a function of large-scale forcing conditions. To provide a quantitative validation of the theoretical distribution, therefore, would require a more sophisticated approach, such as re-scaling the data by the changing mean total mass flux value, $\langle M \rangle$ (assuming a constant value of $\langle m \rangle$), to produce a single probability distribution. The possibility of convective organisation could also, in principle, be incorporated by sub-dividing the data into different regimes, defined by typical values of $\langle m \rangle$. It would seem useful to revisit this analysis, incorporating the influence of the environment, as indicated above.

References

Arakawa, A. and Schubert, W. H., 1974. Interaction of a cumulus cloud ensemble with the large-scale environment, Part 1. *Journal of the Atmospheric Sciences*, **31**, 674–701.

Astin, I. and Latter, B. G., 1998. A case for exponential cloud fields? *Journal of Applied Meteorology*, **37**, 1375–1382.

Buizza, R., Miller, M., and Palmer, T., 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **125**, 2887–2908.

Cahalan, R. F. and Joseph, J. H., 1989. Fractal statistics of cloud fields. *Monthly Weather Review*, **117**, 261–272.

Chandler, 1987. Introduction to modern statistical mechanics. *Oxford University Press.*

Chen, S. S., Houze, R., and Mapes, B. E., 1996. Multiscale variability of deep convection in relation to large-scale circulation in TOGA COARE. *Journal of the Atmospheric Sciences*, **53**, 1380–1409.

Cho, H.-R., 1978. Some statistical properties of a homogeneous and stationary shallow

cumulus cloud field. *Journal of the Atmospheric Sciences.*

Cohen, B. G., 2001. Fluctuations in an ensemble of cumulus clouds. *PhD thesis, University of Reading.*

Cotton, W. R. and Anthes, R. A., 1989. Storm and cloud dynamics. *Academic Press, Inc.*, **44**.

Eagleson, P. S., 1978. Climate, soil, and vegetation 2. the distribution of annual precipitation derived from observed storm sequences. *Water Resources research*, **14**, 713–721.

Emanuel, K. A., 1991. A scheme for representing cumulus convection in large-scale models. *Journal of the Atmospheric Sciences*, **48**, 2313–2335.

Fritsch, J. M., Houze, R. A., adler, R., Bluestein, H., Bosart, L., Brown, J., Carr, F., Davis, C., Johnson, R. H., Junker, N., Kuo, Y.-H., Rutledge, S., Smith, J., Toth, Z., Wilson, J. W., Zipser, E., and Zrnic, D., 1998. Quantitative precipitation forecasting: Report of the eighth prospectus development team, U.S¿ weather research program. *Bulletin of the American Meteorological Society*, **79**, 285–299.

Grabowski, W. W. and Moncrieff, M. W., 2001. Large-scale organization of tropical convection in two-dimensional explicit numerical simulations. *Quarterly Journal of*

the Royal Meteorological Society, **127**, 445–468.

Gregory, D., 1997. The mass flux approach to the parametrization of deep convection. *The physics and parameterization of moist atmospheric convection, NATO ASI series*, **505**, 297–319.

Gregory, D. and Miller, M. J., 1989. A numerical study of the parametrization of deep tropical convection. *Quarterly Journal of the Royal Meteorological Society*, **115**, 1209–1241.

Gupta, V. K. and Waymire, E. C., 1979. A stochastic kinemati study of subsynoptic space-time rainfall. *Water resources research.*

Hozumi, K., Harimaya, T., and Magono, C., 1982. The size distribution of cumulus clouds as a function of cloud amount. *Journal of Meteorological Research Japan.*

Jorgensen, D. P. and Lemone, M. A., 1989. Vertical velocity characteristics of oceanic convection. *Journal of the Atmospheric Sciences.*

Kavvas, M. L. and Delleur, J. W., 1981. A stochastic cluster model of daily rainfall sequences. *Water resources research.*

Landau, L. D. and Lifshitz, E. M., 1968. Statistical physics. *Volume 5 of Course of theoretical physics.*

Lesieur, M., 1997. Turbulence in fluids. *Fluid Mechanics and its Applications, Kluwer Academic Publishers*, **40**.

Lin, J. W. and Neelin, J. D., 2000. Influence of a stochastic moist convective parameterization on tropical climate variability. *Geophysical Research Letters*, **27**, 3691–3694.

Lopez, R. E., 1978. Internal structure and development processes of C-scale aggregates of cumulus clouds. *Monthly Weather Review.*

Lopez, R. E., Blanchard, D. O., Rosenfeld, D., Hiscox, W. L., and Casey, M. J., 1984. Population characteristics, development processes and structure of radar echoes in south Florida. *Monthly Weather Review.*

Moncrieff, M. W. and Klinker, E., 1997. Organized convective systems in the tropical western pacific as a process in general circulation models: A TOGA COARE case-study. *Quarterly Journal of the Royal Meteorological Society*, **123**, 805–827.

Nair, U. S., Weger, R. C., Kuo, K. S., and Welch, R. M., 1998. Clustering, randomness, and regularity in cloud fields 5. the nature of regular cumulus cloud fields. *Journal of*

Geophysical Research, **103**, 11,363–11,380.

Ooyama, K., 1971. A theory on parameterization of cumulus convection. *Journal of the Meteorological Society of Japan*, **49**, 744–756.

Palmer, T., 2001. A nonlinear dynamical perspective on model error: a proposal for nonlocal stochastic-dynamic parametrisation in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, **127**, 279–304.

Peng, L., Sui, C.-H., Lau, K.-M., and Tao, W.-K., 2001. Genesis and evolution of hierarchical cloud clusters in a two-dimensional cumulus-resolving model. *Journal of the Atmospheric Sciences.*

Plank, V. G., 1969. The size distriubtion of cumulus clouds in representative Florida populations. *Journal of Applied Meteorology.*

Ramirez, J. A., Bras, R. L., and Emanuel, K. A., 1990. Stabilization functions of unforced cumulus clouds: their nature and components. *Journal of Geophysical Research*, **95**, 2047–2059.

Randall, D. A. and Huffman, G. J., 1980. A stochastic model of cumulus clumping.

Journal of the Atmospheric Sciences.

Rickenbach, T. M. and Rutledge, S. A., 1998. Convection in TOGA COARE: horizontal scale, morphology and rainfall production. *Journal of the Atmospheric Sciences*, **55**, 2715–2729.

Robe, F. R. and Emanuel, K. A., 1996. Moist convective scaling: Some inferences from three-dimensional cloud ensemble simulations. *Journal of the Atmospheric Sciences*, **53**, 3265–3275.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V., 1987. Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society of London A*, **410**, 269–288.

Sengupta, S. K., Welch, R. M., Navar, M. S., Berendes, T. A., and Chen, D. W., 1990. Cumulus cloud field morphology and spatial patterns derived from high spatial resolution landsat imagery. *Journal of Applied Meteorology*, **29**, 1245–1267.

Stauffer, D. and Ahorony, A., 1992. Introduction to percolation theory. *London: Taylor and Francis.*

Taylor, H. M. and Karlin, S., 1994. An introduction to stochastic modelling. *Academic Press.*

Tiedke, M., 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, **117**, 1779–1800.

Yanai, M., Esbenson, S., and Chu, J.-H., 1973. Determination of bulk properties of tropical cloud clusters from large scale heat and moisture budgets. *Journal of the Atmospheric Sciences*, **30**, 611–627.

Figure Captions

Figure 1 Division of a mass flux interval into two clouds with mass fluxes M_1 and M_2.

Figure 2 Theoretical probability distributions of total mass flux, $p(M)$. (a) $\langle N \rangle = 68$, (b) $\langle N \rangle = 5$. $\langle m \rangle = 10^7\ kgs^{-1}$ for both curves.

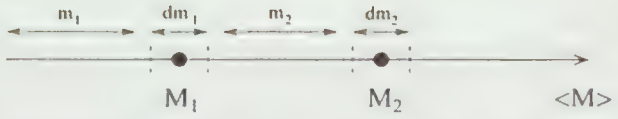
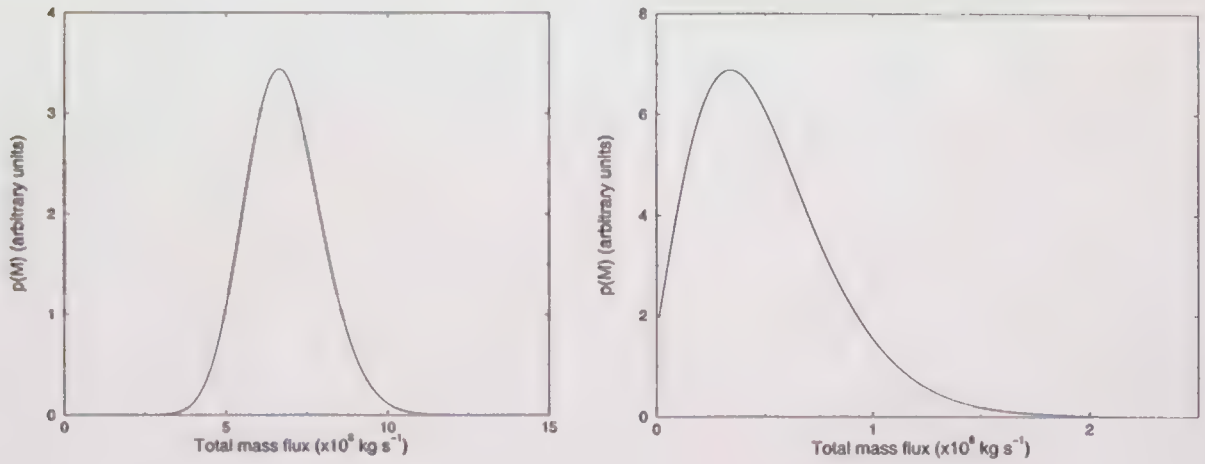


Figure 1: Division of a mass flux interval into two clouds with mass fluxes M_1 and M_2.



(a) (b)

Figure 2: Theoretical probability distributions of total mass flux, $p(M)$. (a) $\langle N \rangle = 68$, (b) $\langle N \rangle = 5$. $\langle m \rangle = 10^7\ kgs^{-1}$ for both curves.

Fluctuations in an equilibrium convective ensemble
Part II: Experimental validation

1 Introduction

In Part I of this investigation, theoretical predictions for the fluctuations of an ensemble of non-interacting convective clouds in equilibrium with the large-scale forcing have been derived by exploiting the analogy between convection and an ideal gas. In particular, the assumption that a total amount of mass flux, M, is distributed randomly amongst the N clouds implies an exponential distribution of individual cloud mass fluxes (analogous to the Boltzmann distribution of molecular energies), *i.e.*:

$$d\overline{n}(m) = \frac{\langle N \rangle}{\langle m \rangle} e^{-m/\langle m \rangle} dm \tag{1}$$

where $d\overline{n}(m)$ is the average number of clouds in the ensemble possessing a mass flux between m and $m + dm$, $\langle N \rangle$ is the ensemble average total number of clouds in the system, and $\langle m \rangle$ the ensemble average mass flux per cloud. All mean quantities are evaluated as horizontal averages at a given height.

By assuming the number of clouds within a fixed region to be random (with mean $\langle N \rangle$), the probability distribution of total mass flux in the ensemble, $p(M)$, has been derived as:

$$p(M) = \left(\frac{\langle N \rangle}{\langle m \rangle} \right)^{1/2} e^{-\langle N \rangle} M^{-1/2} e^{-M/\langle m \rangle} I_1 \left(2 \sqrt{\frac{\langle N \rangle}{\langle m \rangle} M} \right) \tag{2}$$

where M is the total mass flux of the system, and $I_1(x)$ is the modified Bessel function of order 1.

Finally, the normalised variance of the total mass flux distribution was found to be of the form:

$$\frac{\langle(\delta M)^2\rangle}{\langle M\rangle^2} = \frac{2}{\langle N\rangle} \tag{3}$$

In this paper, the validity of this statistical approach is assessed by comparing the theoretical predictions of Equations (1), (2) and (3) against cloud mass flux data obtained from a three-dimensional (3D) Cloud Resolving Model (CRM), in a state of radiative-convective equilibrium.

2 Model formulation and setup

A Cloud Resolving Model is an appropriate choice of numerical tool for this investigation, since it can explicitly resolve cloud-scale dynamics, while also including a parameterisation of sub-grid scale processes. The Met Office Large Eddy Model (LEM) used in the study is summarised briefly in the following section, but is described in more detail by Shutts and Gray (1994) and Tompkins and Craig (1998). The details of the experimental setup are given in Section 2(b).

a. Model formulation

The dynamical core of the CRM is the solution of the anelastic, Boussinesq approximation to the prognostic momentum and thermodynamic equations, given in Shutts and Gray (1994). Liquid water temperature is used as the prognostic thermodynamic variable, since it is reasonably accurately conserved in deep convective ascent (Shutts, 1991). The advection scheme is the Total Variance Diminishing (TVD) scheme of Leonard (1991), which is capable of advecting sharp gradients without overshoots.

Parameterisation of sub-grid scale turbulence is accomplished using a first order Smagorinsky-Lilly closure (Mason, 1989), modified to include a dependence on the pointwise moist Richardson number (MacVean and Mason, 1990). Surface fluxes are parameterised using similarity theory. The microphysical parameterisation used is the 3 phase ice scheme of Swann (1994) and Brown and Swann (1997), with prognostic equations for 6 microphysical variables: mass mixing ratios of liquid cloud water, rain drops, ice crystals, snow crystals, graupel and also the number concentration of ice crystals.

The domain size used is $128 \times 128 \times 21$ km^3, with horizontal resolution of 2 km and a stretched grid in the vertical (100 m in boundary layer; 500 m in upper troposphere). A Newtonian damping layer is applied over a layer of 3 km, from a height of 16.75 km, with a damping coefficient of 0.001 s^{-1}. The lateral boundary conditions are bi-periodic.

b. Experimental setup

The CRM is set up to simulate a region of the tropical atmosphere in radiative-convective equilibrium. The radiative forcing is provided via a constant, horizontally homogeneous cooling rate, imposed up to a height of 400 mb (decreasing linearly to zero between 400 and 200 mb), and applied over an ocean surface at a fixed, uniform sea surface temperature of 300 K. No mean wind is imposed throughout the model domain, and the atmosphere is assumed to be non-rotating. By deliberately omitting processes such as cloud-radiative interactions, wind shear and background rotation that are known to cause convective organisation (see, for example, Tompkins and Craig, 1998; Thorpe *et al.*, 1982; Vallis *et al.*, 1997), the intention is to generate an idealised equilibrium convective ensemble most suited to comparison with the statistical theory derived in Part I. The model is run into a full radiative-convective equilibrium state (reached after

10-20 days of simulation, depending on forcing), and the variability of the convective ensemble within this equilibrium period is compared with theory.

Space-time averages over the full horizontal extent of the domain at a height of 2.4 km (corresponding to the base of the deep convective layer), and over a period of several hours within the equilibrium state, are used to approximate the theoretical ensemble average quantities in Equations (1), (2) and (3). To increase the robustness of these statistics, an unusually high cooling rate of $-16\ Kday^{-1}$ is used in an attempt to increase the mean number of clouds in the domain. The sensitivity of the results to the magnitude of the forcing is explained in Section 4.

3 Mass flux per cloud

The CRM is first used to verify the existence of the predicted exponential distribution of mass flux per cloud, given in Equation (1). The mass flux of each cloud within the CRM equilibrium state is diagnosed using a routine which scans through the domain and locates cloudy grid-points; adjacent cloudy points are considered to be part of the same cloud and the size, σ_i, is recorded. Within each of these clouds, an average vertical velocity, $\overline{w_i}$, is computed and the mass flux per cloud, $m_i = \rho\sigma_i\overline{w_i}$, is calculated.

Two different definitions are introduced to identify a 'cloudy' grid-point. Firstly, the standard velocity criterion, $w > 1$ ($w < -1$) ms^{-1} is used to define up(down)-draught points separately (LeMone and Zipser, 1980; Zipser and LeMone, 1980). In the second definition, a cloudy point is identified by the presence of liquid or ice cloud water (*i.e.*, $Q_{CLD} > 0$). In general, these two criteria will produce slightly different results, and their relative accuracy in re-producing the theoretical Boltzmann distribution of

individual cloud mass fluxes makes an additional, interesting comparison.

Figures 1(a) and 1(b) show the mass flux histograms from the $-16\ Kday^{-1}$ simulation equilibrium state, computed using the two separate definitions of $Q_{CLD} > 0$ and $w > 1$ ms^{-1}. Immediately noticeable from Figure 1(a) is that the Q_{CLD} definition has produced a mass flux distribution containing clouds with both net upward and downward motions. Physically, the negative mass flux clouds correspond to convective cells near the end of their life-cycle, where downdraught motions dominate over the updraught component. Nevertheless, this co-existence of both negative and positive clouds is not represented in the theoretical Boltzmann distribution of Equation (1), and a direct comparison between experiment and theory is potentially problematic. One solution to this problem would be to consider the *time-integrated* mass flux per cloud, as suggested in Part I. However, while consideration of such a quantity would effectively remove the negative mass flux points (the net mass flux of a cloud will be positive when integrated over its life-cycle), thus allowing a more realistic comparison to the idealised gas system, this technique would be difficult to implement. A second possibility is to assume a separate Boltzmann distribution for both upward and downward moving clouds; removal of the small number of $w < 0\ ms^{-1}$ points from the experimental data would then allow a direct comparison with the theoretical updraught distribution. This rather simpler solution is the one employed in the remainder of this study, although there is only enough data to test the updraught distribution. The problem of negative mass fluxes is necessarily avoided in the $w > 1\ ms^{-1}$ distribution of Figure 1(b). Here, however, the rather arbitrary, unphysical cut-off at $1\ ms^{-1}$ implies that the weakest clouds predicted by the Boltzmann distribution will be unrepresented by this definition.

Despite these obvious discrepancies from theory, both experimentally derived mass flux distributions in Figure 1 appear to be exponential, as predicted from statistical mechanics. A quantitative comparison with theory is provided by re-plotting the mass flux histograms of Figure 1 with a logarithmic scaling of the y-axes, and by fitting a linear regression through the data. The slope of the regression line should give an estimate of the mean mass flux per cloud, $\langle m \rangle$, while the intercept is related to the mean number of clouds, $\langle N \rangle$. Table 1 compares the fitted values of $\langle m \rangle$ and $\langle N \rangle$ with those diagnosed from the model and, as can be seen, the agreement is extremely good.

4 Total mass flux and variance

The mass flux within a fixed region of the CRM domain is compared directly to the theoretical total mass flux probability distribution of Equation (2). Figures 3(a) and (b) show histograms of the total mass flux within the full 128×128 km^2 model domain and over a 32×16 km^2 sub-area respectively. Qualitatively at least, the shapes of these curves are similar to the theoretical form (*c.f.* Figure 2 of Part I), with a maximum in probability around the mean value of mass flux, and variance and skewness which decreases as the number of clouds increases.

A quantitative comparison with the theoretically derived mass flux probability distribution is provided by applying the analytical form of Equation (2) as a non-linear curve-fit, with $\langle N \rangle$ and $\langle m \rangle$ as free parameters, to the full 128×128 km^2 CRM data (full line in Figure 3(a)). A second curve-fit is also drawn, where now the mean cloud number, $\langle N \rangle$, is constrained to match the simulated mean value (dashed line). Table 2 contains a summary of the values of $\langle N \rangle$ and $\langle m \rangle$ obtained as best-fit parameters from

these two curve-fits, as well as the associated mass flux variance, $\langle(\delta M)^2\rangle$, calculated from Equation (3). The values of $\langle N\rangle$, $\langle m\rangle$ and $\langle(\delta M)^2\rangle$ diagnosed directly from the CRM are also shown for comparison. From this it can be seen that the theoretical form of $p(M)$ predicts a variance which is much larger than the mass flux variability which occurs in practice.

This discrepancy between the theoretical and model variances can be seen more quantitatively in Figure 4, which shows the observed normalised mass flux variance, computed for a range of different sized sub-areas within the domain of the $-16\ Kday^{-1}$ simulation, and plotted directly as a function of $\frac{1}{\langle N\rangle}$ (black points). The corresponding theoretical expression computed from Equation (3) is also shown (dashed line), and a significant and consistent under-estimate of the predicted variance by the CRM simulation is clearly observed.

In deriving the theoretical form of $p(M)$ (Equation 2), the assumptions invoked were an exponential distribution of cloud mass fluxes, and a random positioning of those clouds within a given region. Since the exponential cloud mass flux distribution has already been validated, the lack of variance in the CRM simulation might indicate that the clouds are distributed more regularly than randomly throughout the model domain. Such regularity has been seen previously in both observed and simulated cloud fields (Ramirez and Bras, 1990; Nair *et al.*, 1998), and Ramirez *et al.* (1990) proposed an 'inhibition hypothesis' to explain this behaviour, whereby the subsidence region surrounding each cloud acts locally to inhibit convection, causing a 'repulsive' interaction between neighbouring clouds.

The local effect of one cloud on its neighbours can be determined for the CRM data by computing the mean cloud number in annular regions surrounding each cloud, as a

function of distance from cloud centre (Figure 5). This statistic is similar to that used in Nair *et al.* (1998), although here the positioning and width of the annular regions are not rescaled using the effective mean radii of clouds. The mean cloud number density as a function of distance is shown in Figure 6, and shows a reduction on average in the density of clouds within a 15 km radius of each cloud. At sufficiently large distances from a cloud, this inhibition effect becomes negligible, and the mean cloud density approaches the value expected from a purely random distribution of cloud centres (dashed line). It can be shown that the peak observed in the cloud number density at very small distances from the cloud ($< 5\ km$), as well as some of the subsequent oscillations, are artifacts of the finite resolution of the CRM, and should not be interpreted physically.

Despite the fact that the simulated cloud field is not random, as assumed in the derivation of the theoretical total mass flux distribution, Figure 4 shows clearly that these local cloud interactions do not change the form of the mass flux variability, but simply re-scales its magnitude. A simple stochastic point process model (Strauss, 1975), in which points closer a critical distance are suppressed (analogous to the inhibiting effect of nearby clouds), shows a qualitatively similar decrease in the variability (not shown). This suggests that the theoretical expressions for mass flux variability could be modified to take account of these cloud interactions simply by introducing a single constant re-scaling factor in Equations (2) and (3). The dot-dashed line in Figure 4 shows the result of applying an appropriate correction factor (of around 0.58) to the theoretical variance curve (dashed line), and the close match with the CRM data is now clearly seen. The reason for the remaining small off-set is unclear, but may be a finite domain size effect. The same re-scaling factor is also applied to the $32 \times 16\ km^2$ mass

flux distribution (dot-dashed line in Figure 3(b), and the modified theoretical curve is observed to fit the data extremely well.

Physically, it would be expected that the inhibiting effect of clouds would be more noticeable in regions where the clouds are closely packed, and hence that the relative modulation of the mass flux variability should be greatest when the mean cloud spacing is small. Since the mean number of clouds in the equilibrium ensemble (and hence the cloud spacing) is a direct function of the large-scale forcing (Cohen, 2001), CRM model simulations with different radiative cooling rates can be used to produce equilibrium states with varying mean cloud densities. Figure 7 shows the mass flux histogram for a $-4\ Kday^{-1}$ simulation, together with the constrained and unconstrained forms of the theoretical curve-fit (*c.f.* Figure 3(a)), and it is clear that the unmodified statistical theory predictions are a much better approximation in this lower forcing case. In Figure 8, the ratio of predicted to observed variance for four simulations with cooling rates of -16, -12, -8, -4 and -2 $Kday^{-1}$ is shown as a function of mean cloud spacing. A systematic decrease in this ratio as a function of the applied forcing is clearly exhibited. This, then, shows that the value of the constant re-scaling factor for the mass flux variability should be a function of large-scale forcing. It is significant that at low enough forcings (large mean cloud spacings), the magnitude of the re-scaling parameter appears to tend towards unity, which corresponds to the low-density, ideal gas limit, where the theory would be expected to apply most accurately. But it appears that his ideal limit is not approached for most forcings of meteorological interest.

At present, the value of the variance correction term must be determined empirically for each situation but, in principle, a physical model specifying the inhibition function

could be developed, which would allow quantitative, analytical predictions of the magnitude of the variability re-scaling parameter to be made. Such an inhibition model would need to account for the fact that the clouds have different sizes and intensities and will likely have different inhibition strengths and different susceptibilities to inhibition.

5 Conclusions and implications

The CRM simulations presented in this paper have provided considerable validation of the statistical theory of convective fluctuations outlined in Part I of this study, but with one significant discrepancy. In particular, the exponential probability distribution of individual cloud mass fluxes recovered from the CRM data is exactly in line with the theoretical prediction. However, a similar analysis of the total mass flux distributions highlights that the mass flux variability in these simulations is significantly less than would be predicted for an ensemble of non-interacting clouds. Further analysis has shown that clouds act to suppress convection within a localised region (similar to the inhibition mechanism hypothesised by Ramirez *et al.*, 1990), and these local interactions are sufficient to explain the observed discrepancy from theory.

Interestingly, it has been found that the theoretical total mass flux probability distribution and associated variance expression can be satisfactorily modified to take account of this inhibition effect by introducing a simple, constant re-scaling factor, which is a function of the mean cloud spacing. This correction factor allows the exact form of the mass flux fluctuations around equilibrium to be quantitatively predicted. The development of a physical model of the inhibition process is clearly desirable, since the complete probability density function could then be derived using the methods of Part I.

References

Brown, P. R. A. and Swann, H. A., 1997. Evaluation of key microphysical parameters in three-dimensional cloud-model simulations using aircraft and multiparameter radar data. *Quarterly Journal of the Royal Meteorological Society*, **123**, 2245–2275.

Cohen, B. G., 2001. Fluctuations in an ensemble of cumulus clouds. *PhD thesis, University of Reading.*

LeMone, M. A. and Zipser, E. J., 1980. Cumulonimbus vertical velocity events in GATE. Part I: diameter, intensity and mass flux. *Journal of the Atmospheric Sciences*, **37**, 2444–2457.

Leonard, B. P., 1991. The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Comput. Methods Appl. Mech. Eng.*, **19**, 17–74.

MacVean, M. K. and Mason, P. J., 1990. Cloud-top entrainment instability through small-scale mixing and its parameterization in numerical models. *Journal of the Atmospheric Sciences*, **47**, 1012–1030.

Mason, P. J., 1989. Large-eddy simulation of the convective atmospheric boundary layer. *Journal of the Atmospheric Sciences*, **46**, 1492–1516.

Nair, U. S., Weger, R. C., Kuo, K. S., and Welch, R. M., 1998. Clustering, randomness, and regularity in cloud fields 5. the nature of regular cumulus cloud fields. *Journal of Geophysical Research*, **103**, 11,363–11,380.

Ramirez, J. A. and Bras, R. L., 1990. Clustered or regular cumulus cloud fields: the statistical character of observed and simulated cloud fields. *Journal of Geophysical Research*, **95**, 2035–2045.

Ramirez, J. A., Bras, R. L., and Emanuel, K. A., 1990. Stabilization functions of unforced cumulus clouds: their nature and components. *Journal of Geophysical Research*, **95**, 2047–2059.

Shutts, G. J., 1991. Liquid water static energy - a moist thermodynamic variable for use in numerical models. *Met Office, JCMM mesoscale newsletter*, **No. 3**, 36–41.

Shutts, G. J. and Gray, M. E. B., 1994. A numerical modelling study of the geostrophic adjustment process following deep convection. *Quarterly Journal of the Royal Meteorological Society*, **120**, 1145–1178.

Strauss, D. J., 1975. A model for clustering. *Biometrika*, **62**, 467–475.

Swann, H., 1994. Cloud microphysical processes - a description of the parametrization used in the large eddy model. *Met Office, JCMM Internal Report*, No. 10.

Thorpe, A. J., Miller, M. J., and Moncrieff, M. W., 1982. Two-dimensional convection in non-constant shear: a model of mid-latitude squall lines. *Quarterly Journal of the Royal Meteorological Society*, **108**, 739–762.

Tompkins, A. M. and Craig, G. C., 1998. Radiative-convective equilibrium in a three-dimensional cloud-ensemble model. *Quarterly Journal of the Royal Meteorological Society*, **124**, 2073–2097.

Vallis, G. K., Shutts, G. J., and Gray, M. E. B., 1997. Balanced mesoscale motion and stratified turbulence forced by convection. *Quarterly Journal of the Royal Meteorological Society*, **123**, 1621–1652.

Zipser, E. J. and LeMone, M. A., 1980. Cumulonimbus vertical velocity events in GATE. Part II: synthesis and model core structure. *Journal of the Atmospheric Sciences*, **37**, 2458–2469.

List of Tables

Table 1 Values of $\langle m \rangle$ and $\langle N \rangle$, diagnosed from the $-16\ K day^{-1}$ simulation of Figure 2, and calculated from the best-fit regression for both mass flux definitions.

Table 2 $\langle N \rangle$, $\langle m \rangle$ and $\langle (\delta M)^2 \rangle^{1/2}$ diagnosed from the full domain, $-16\ K day^{-1}$ simulation of Figure 3(a), and from two different theoretical curve-fits.

Figure Captions

Figure 1 Histograms of mass flux per cloud at a height of 2.4 km from the -16 $Kday^{-1}$ CRM simulation, using mass flux criteria (a) $Q_{CLD} > 0$, (b) $w > 1ms^{-1}$.

Figure 2 As for Figure 1, but with logarithmic scaling of the y-axes. The linear regressions are also plotted (dashed lines).

Figure 3 Probability distributions of total mass flux, $p(M)$, at a height of 2.4 km from the -16 $Kday^{-1}$ simulation. (a) 128×128 km^2 domain, with unconstrained curve-fit (full line); constrained curve-fit (dashed). (b) 32×16 km^2 sub-area, with modified theoretical prediction (dot-dashed line).

Figure 4 The normalised standard deviation of total mass flux as a function of the square root of inverse cloud number, from the -16 K day^{-1} simulation. Unmodified theoretical prediction (dashed line); modified by constant re-scaling factor (dot-dashed).

Figure 5 Schematic of annular regions used to compute mean cloud number density as a function of radius, D, from centre of a cloud for CRM velocity field.

Figure 6 Cloud number density as a function of distance from cloud centre. Dashed line is theoretical prediction for Poisson distribution of clouds with same mean cloud density.

Figure 7 Same as Figure 3(a), but for the -4 $Kday^{-1}$ simulation.

Figure 8 Ratio of theoretical to modelled variance of total mass flux as a function of mean cloud spacing from the -16, -12, -8, -4 and -2 $Kday^{-1}$ CRM simulations.

	$\langle m \rangle$ ($\times 10^7$ kgs^{-1})		$\langle N \rangle$	
Mass flux Definition	CRM	Regression	CRM	Regression
$w > 1 \mathrm{ms}^{-1}$	3.16	3.19	68.35	67.87
$Q_{CLD} > 0$	3.15	3.16	63.59	62.92

Table 1: Values of $\langle m \rangle$ and $\langle N \rangle$, diagnosed from the -16 $Kday^{-1}$ simulation of Figure 2, and calculated from the best-fit regression for both mass flux definitions.

	Theoretical curve-fit		CRM
	$\langle N \rangle$, $\langle m \rangle$ free	$\langle N \rangle$ constrained	
$\langle N \rangle$	246.8	68.35	68.35
$\langle m \rangle$ ($\times 10^7$ kgs^{-1})	0.92	3.34	3.16
$\langle (\delta M)^2 \rangle^{1/2}$ ($\times 10^8$ kgs^{-1})	2.05	3.91	2.03

Table 2: $\langle N \rangle$, $\langle m \rangle$ and $\langle (\delta M)^2 \rangle^{1/2}$ diagnosed from the full domain, -16 $Kday^{-1}$ simulation of Figure 3(a), and from two different theoretical curve-fits.

Dashed line is best-fit power law, of form $y = 336x^{-1.8} + 0.94$.

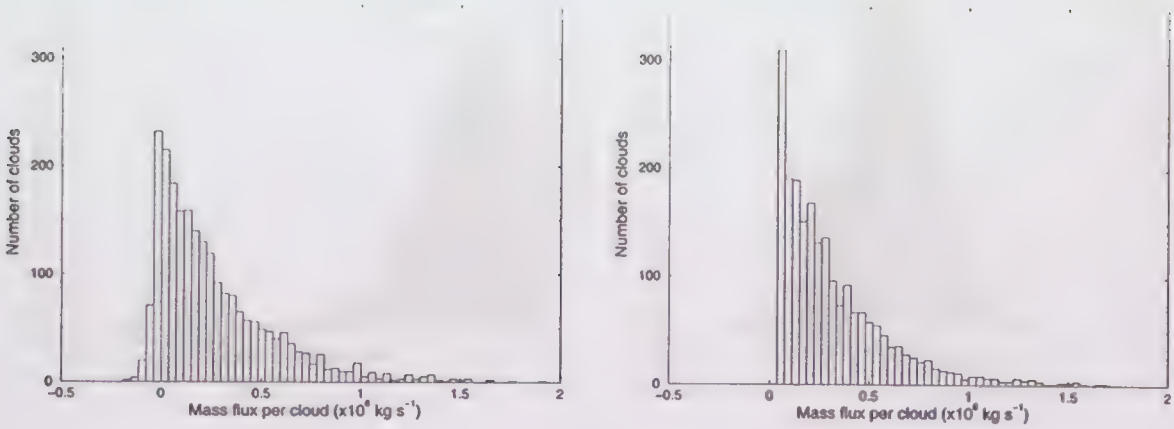


Figure 1: Histograms of mass flux per cloud at a height of 2.4 km from the $-16\ Kday^{-1}$ CRM simulation, using mass flux criteria (a) $Q_{CLD} > 0$, (b) $w > 1ms^{-1}$.

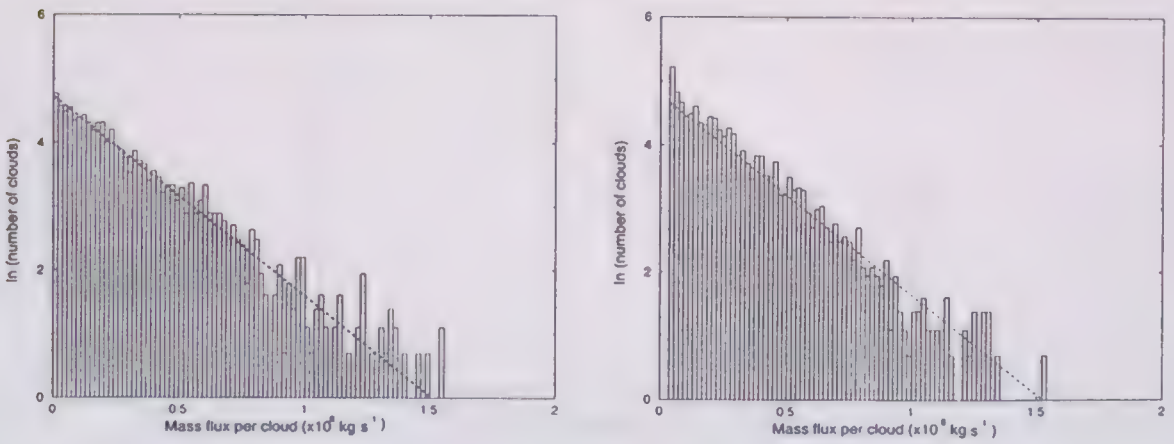


Figure 2: As for Figure 1, but with logarithmic scaling of the y-axes. The linear regressions are also plotted (dashed lines).

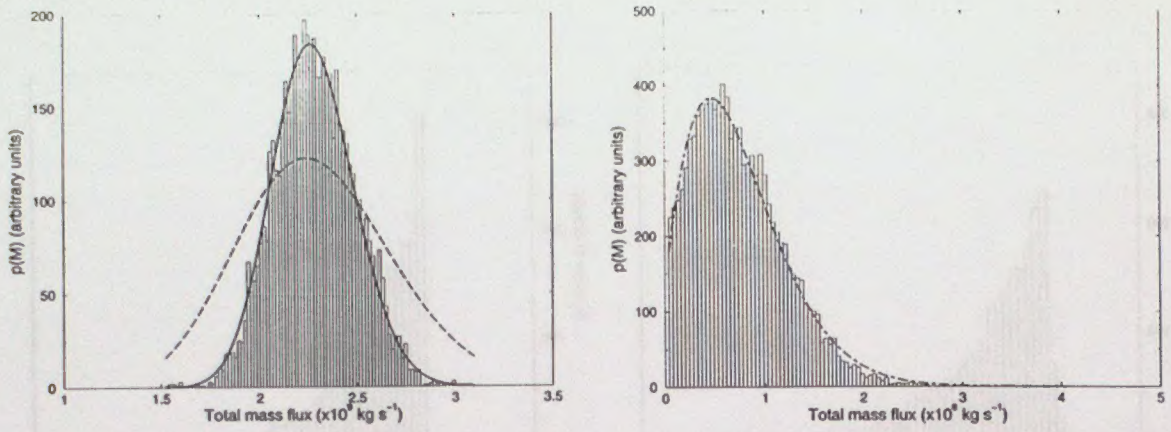


Figure 3: Probability distributions of total mass flux, $p(M)$, at a height of 2.4 km from the -16 $Kday^{-1}$ simulation. (a) 128×128 km^2 domain, with unconstrained curve-fit (full line); constrained curve-fit (dashed). (b) 32×16 km^2 sub-area, with modified theoretical prediction (dot-dashed line).

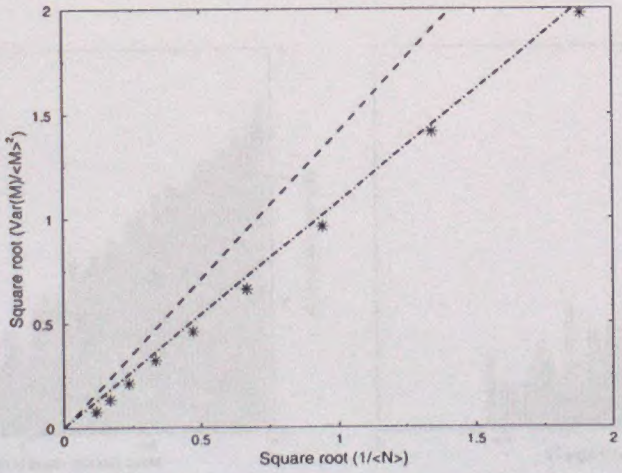


Figure 4: The normalised standard deviation of total mass flux as a function of the square root of inverse cloud number, from the -16 K day^{-1} simulation. Unmodified theoretical prediction (dashed line); modified by constant re-scaling factor (dot-dashed).

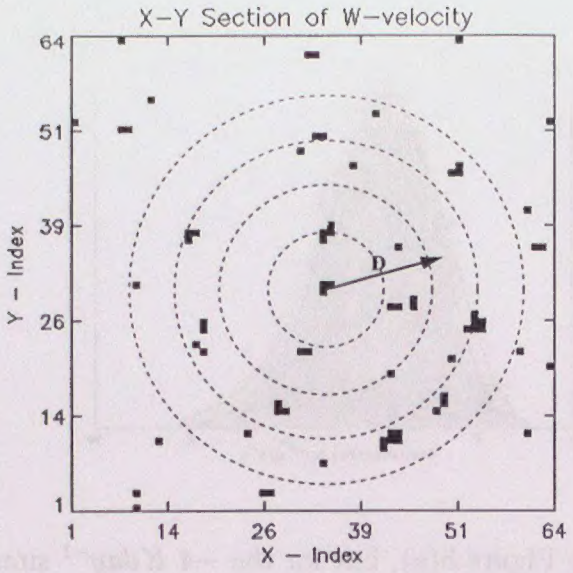


Figure 5: Schematic of annular regions used to compute mean cloud number density as a function of radius, D, from centre of a cloud for CRM velocity field.

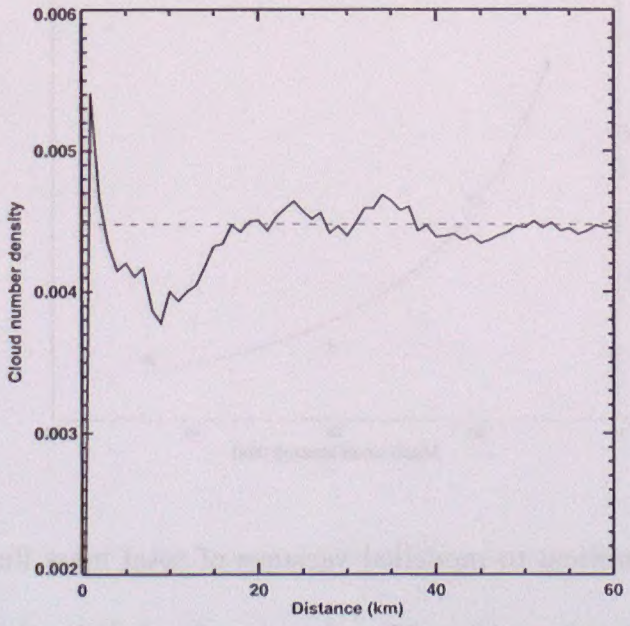


Figure 6: Cloud number density as a function of distance from cloud centre. Dashed line is theoretical prediction for Poisson distribution of clouds with same mean cloud density.

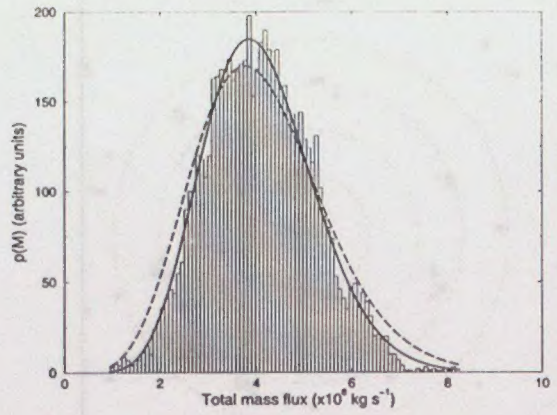


Figure 7: Same as Figure 3(a), but for the $-4\ Kday^{-1}$ simulation.

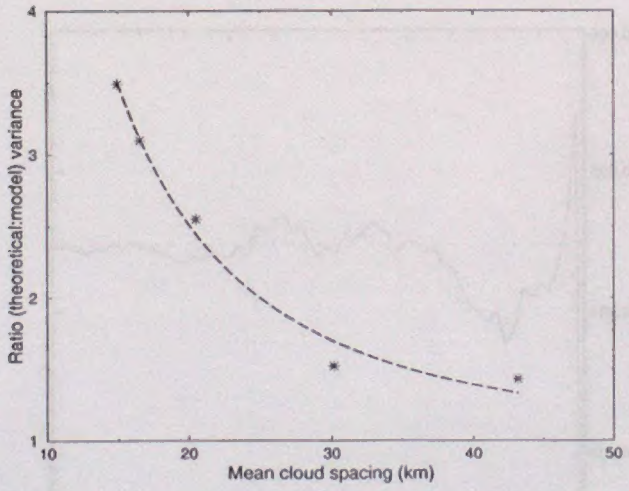


Figure 8: Ratio of theoretical to modelled variance of total mass flux as a function of mean cloud spacing from the -16, -12, -8, -4 and $-2\ Kday^{-1}$ CRM simulations. Dashed line is best-fit power law, of form $y = 336x^{-1.8} + 0.94$.